Animal Pose Estimation: Cross-Species Data and Keypoint Schemes

Shaan Chanchani, Claire Kim, Josh Mansky, Zian Pan, Medhashree Parhy, Parth Thakre

With thanks to Haoyu Chen and Dr. Amy Reibman

Overview

- Improve robustness and accuracy of antelope keypoint estimation models with the downstream task of improving pose estimation to aid in species preservation
- We utilize two different approaches to investigate
  1. Focusing the training data to more similar species, which improves model generalization
  2. Re-labeling the data with two custom keypoint definitions, which improves consistency of data as compared to AP10k annotations
- Using RTMPose model and AP10k data

Takeaways

- Using similar species with respect to the test data has a high likelihood of performing better than random/large dataset approaches.
- The definition of similar species makes a difference, as some methods of quantifying similarity can yield different levels of performance
- Using a keypoint labeling scheme that can ensure better consistency across labelers and better defined keypoints, can improve model performance

Next Steps

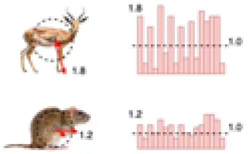
- Adjust model training process (made for 17 keypoints) to fit ***Biological*** keypoint definition (21 keypoints)
  - This can let us further test our theory of keypoint definition playing a large role in model performance
- Improve pose estimation models performance on in-the-wild antelopes
  - Apply ***Visible*** and ***Biological*** definitions to these images to get robust annotations
  - Include camera-trapped similar species to further focus the data

Data Focusing Based on Similar Species

Does a model perform better when it's trained on similar species?

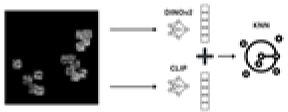
Key Question: How can we define the list of the most visibly similar species?

1. Centroid Approach:



1. Limb Ratios:

1. Dino+SD V2:



Metric	AP-10k Redistributed	Reduced AP-10k Redistributed	Centroid	Limb Ratios	Dino+SD V2	Human Ranking	Taxonomic Species
Average Precision	0.818	0.752	0.825	0.794	0.804	0.821	0.774

Remarks:

- When normalized by size, the similar-species focused datasets achieve better results than the AP-10K dataset.

Re-labeling Data

Does a model perform better when its training data is more well-defined and consistent?

Problem:

The existing AP10k (red in figure below) annotations have inconsistent keypoint definitions across images, which may be hindering model performance

Keypoint Definitions:

We create two keypoint definitions, ***Visible*** - which allows labelers to pick more apparent points (light blue in figure) and processes the midpoint for those points (blue in figure), and ***Biological*** - based on anatomically accurate points, to see how model performance varies based on keypoint definition

Red -AP10k, Blue - Visible Processed, Light Blue - Visible Unprocessed

Metric	Visible Keypoints	AP-10k
Average Precision	0.725	0.692

Remarks:

- The ***Visible*** definition outperformed the AP10k labeled data
- Our created definition enables more consistent labeling at accurate points on the antelope, which improves model performance